

AMENDMENTS

Amendments to the Specification:

In response to the Official Action and in accordance with 37 CFR 1.121(c), please enter the following rewritten paragraphs from the specification of the instant application.

Please amend paragraphs [0016], [0040], [0046], [0059], [0061], [0062], [0063] and [0064] as is indicated below:

[0016] In another preferred embodiment of the alignment process, the instant invention uses a more elaborate method of locating change point markers in the audio work. In brief, in this preferred embodiment multiple criteria are used to locate markers in the audio data. Of course, that should yield a larger number of potential markers against which to compare the breaks/discontinuities in the video data. Preferably, though, the audio markers will be matched against the video data according to an order specified by the user. For example, the algorithm might attempt, first, to match markers produced by a “volume level” algorithm. Then, if none of the markers that were obtained by the “volume” method is are satisfactory, the algorithm could use markers produced by a beat detection algorithm, etc. Needless to say, because of the increased complexity of this embodiment, additionally additional computer processing power may be needed.

[0040] As a next preferred step, the user will signal to the computer that an analysis of the audio data should be performed (step 530). One product of this analysis (discussed in more detail below in connection with Figure 6) is the selection and posting of the audio

markers (e.g., M1 to M7 of Figures 3 and 4). Next, the user will typically request that the data in the video track be analysed for breaks and/or the user might manually determine where such breaks should appear in the video work (e.g., markers T1, T2, and T3 in Figure 3 and 4). Obvious locations for breaks include, by way of example, junctions between time-adjacent video clips, locations within a video clip where substantial changes in illumination level occur within the space of a few frames, locations where there are substantial changes in the quantity of on-screen motion or activity, etc. Any of the foregoing might be automatically designated as a location where a video marker could be placed. Preferably the user will thereafter determine which transition effects will be applied at each video marker location, with these effects typically being selected from a menu of standard transitions, but such selections could obviously be made by the program instead. Additionally it should be noted that the algorithms for identifying scene cuts in an existing video work depends depend on a number of different factors and parameters. Thus, it is not unexpected that the user might wish to review the automatic placement of the scene cuts/transitions and alter such by moving, deleting, adding to, etc. them. As a consequence, it is anticipated that such editing options will typically be provided to the user. Finally, it is also anticipated that a user might have manually flagged certain locations within the video work for transitions and/or alignment. If that is the case, it is preferred that the instant algorithm would leave those marks undisturbed.

[0046] Next, the audio clip or clips will preferably be scanned for additional sonic variations that could serve as marker locations (step 625). Preferably, this process will be initiated by the user although it could certainly take place automatically. In either case, it is

anticipated that several different detection and scanning schemes will be employed and they will preferably be carried out sequentially on the audio data. For example, some preferred analyses include is an examination of the musical work or clip for entry into/exit from a chorus or verse, changes in musical key or pitch, changes in strophe or other musical phrasing, changes in frequency content/centre frequency (e.g., possibly representing the occurrence of solos by different musical instruments), timbre, lead instrument, volume, etc. Those of ordinary skill in the art will recognize that these properties are just a few of the many that could potentially be used.

[0059] The following example is offered to illustrate further the operation of one preferred embodiment of the instant invention, wherein the video clip lengths are shortened or lengthened to cause them to match the markers in the musical work. As is generally indicated in Figure 9, two audio markers 930 and 935 have been detected within the audio track 910, one at 20 seconds and another at 30 seconds respectively. The video track 920 will be assumed to contain at least three video clips. In this example, video 1 has a displayed length of 17 seconds and an unedited length of 25 seconds. Video clip 2 has a displayed length of 10 seconds and an unedited length of 15 seconds. Finally, video clip 3 has a displayed length of 18 seconds and an unedited length of 20 seconds. Said another way, each of the video clips has been edited to shorten its visible playtime from its original (unedited) length to the length indicated in the figure.

[0061] Next, an inquiry will preferably made as to whether it would be possible to lengthen video clip 1 by the requisite amount and correspondingly shorten the length of video

clip 2 so as to cause the transition between the first two video clips to coincide with audio markers 930 and to leave the transition between video clips 2 and 3 unremoved. Thus, it will preferably next be determined as to whether under the current rule structure it would be possible to shorten video clip 2 by three seconds. If so, that operation (i.e., lengthening clip 1 and shortening clip 2) will preferably be performed. Note that video clip 2 might be shortened either by removing time from the start ~~of~~ or end of the clip. For purposes of illustration, it will be assumed to be removed from the end.

[0062] Then, as a next step, the algorithm will preferably attempt to synchronize the transition between video clips 2 and 3 with audio marker 935. Note that, after the previous steps, video clip 2 has a displayed length of 7 seconds, which means that if needed it could be extended by as much as 8 seconds. Next, the time-difference between the marker 935 at 30 seconds and the closest video transition (i.e., the one between video clips 2 and 3) will preferably be calculated to be three seconds (30 seconds – (20 seconds + 7 seconds)). Thus, one preferred method of synchronizing the transition between clips 2 and 3 with the marker **935** is carried out by extending the displayed length of video clip 2 by three seconds. Of course, that will only be possible if additional video footage is available (which it is). Recall, that video clip 2 could be lengthened by as much as eight seconds (i.e., the current displayed length is 7 seconds out of a total unedited length of 15 seconds). Thus, by adding three additional seconds to video 2 (preferably at its end where the same amount of video footage was removed previously) the transition between video clips 2 and 3 may be moved to the 30 second time point where it will coincide with audio marker 2. Finally, it is ~~preferably~~ preferred that video clip 3 be shortened by three seconds, so that its end point does not move, preferably at its ending although that choice

could be left to the user. Of course, all of the foregoing was done under the assumption that none of the operations would cause any of the video clips involved to be shortened or lengthened beyond a permitted value.

[0063] Note that it is anticipated that the preferred method of extending and shortening a video clip will be to use conventional video editing techniques to make more or less of a video clip viewable within the video work. However, those of ordinary skill in the art will recall, as has been discussed previously, that alternatively (or perhaps in conjunction with the previous steps) a video effect could be chosen that overlays less or more of the selected video clips, thereby effectively obscuring the actual transition point between the two clips and making it appear as though the transition coincides with the audio marker. As an example, and returning to the scenario discussed in the previous paragraphs in connection with Figure 9, rather than actually shortening or lengthening video clips 1 and 2, a multi-second transition (e.g., the transition might be a long fade-to-black followed by an abrupt return to full brightness at the following video frame) might be applied which would overlay the start of video clip 2 (and possibly the ending of video clip 1) and end at the time point that corresponds to audio marker 930, thereby making video clip 2 fully viewable again beginning at 20 seconds into the video work. Thus, for purposes of the instant invention when the viewable portion of a clip is described as being “shortened” that term should be understood to include shortening of its viewable portion by displaying fewer frames as well as shortening it by obscuring a portion of that clip with a video transition effect. Similarly, when the viewable portion of a clip is “lengthened” that term should be understood to include making additional frames visible using conventional video editing methods or decreasing the coverage of transition effects, thereby uncovering more of the clip. Further, it should be

remembered that in those instances where a clip is to be lengthened beyond its unedited length, there are any number of conventional methods of increasing the displayed length of a video clip even if additional video frames are not available.

[0064] Finally, it is certainly possible that a user might not object to having the instant program relocate video clips in time in order to synchronize one or more video transitions with musical markers. That is, in still another preferred embodiment the instant invention might optionally operate as follows. Assume for purposes of illustration, that a video marker has been selected at the junction between two time-adjacent video clips. Suppose further, that it is possible that ~~by~~ shortening the viewed length of the leading clip ~~to bring~~ brings the junction into alignment with the selected audio marker. In this embodiment, in contrast to what was done previously, the first clip would be shortened to cause the end of this clip to at least approximately coincide with the audio marker. Then, the clip that follows would be slid in time to cause its start time to once again abut the ending of the now-shortened clip. Preferably, the clips that follow would be similarly moved, so that the net result would be – unless other adjustments were made – a corresponding shortening in the play time of the video work. This would, of course, have the benefit of leaving the second/later clip (and the clips that follow) completely unmodified which might be desirable in some circumstances. Of course, it should be clear that this idea could readily be incorporated into the preferred embodiments discussed previously. That is, some combination of shortening/sliding and modification of the transition parameters could certainly be used. In most circumstances, this will preferably be left to the desires of the user.